International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Depression Detection Using Machine learning

Ms. Sonal Kadam[1], Ms. Vaibhavi Thorat[2], Dr.Harshita Vachhani[3]

*[*1,2 PG] Computer Science, Pratibha College of Commerce and Computer Studies, Chinchwad, Pune,* Maharashtra, India.

*[*3]Professor, Co Computer Science, Pratibha College of Commerce and Computer Studies, Chinchwad, Pune,*
Maharashtra, India.

Abstract - According to the World Health organization, by2030depression will be the second leading cause of disable Depression is state of Mental illness. It is characterized by long lasting feeling of sadness or despair. Most patients with depression do not complain that they are depressed. If a person is sad for really long time, then that person can be considered as depressed person. such person needs to take the proper diagnosis with the help of psychiatrist. But this people did not like visit the psychiatrist to check that if they are depressed or not because they are fear people judgements. In recent years machine learning (ML) Has emerged as a powerful tool for the automatic detection of depression by analyzing patterns in behavioral and physiological data. this research explore, the application of ML algorithms to identify depressive symptoms using data derived from source such as social media post, speech signals and structure questionaries. We evaluate multiple classify including support vector machines. (SVM), Random Forest and deep learning models to determine their effectiveness in detecting depression. So, this project can help such people to check on their own if they are depressed or not.

Keywords: Depression Detection, Facial Expressions, Social media Analysis, Support vector machines (SVM), Natural language Processing (NLP), Speech Analysis, Machine Learning

1.INTRODUCTION

One of the increasing mental health problems affecting many of the younger generation is depression. Depression is caused by work culture, demanding lifestyles, emotional imbalances family issues, and social life. Depression is becoming a common and serious illness that negatively impacts people's daily lives. Such mental states usually lead to grief, loss of interest in things and work, and can lead to suicide. It interferes with natural functioning both at work
and at home. The main symptoms of depression are changes in eating and sleeping habits, low energy, difficulty concentrating, anxiety, hopelessness. Thoughts of self-harm, and self-harm. Main results include weight loss, heart disease, and inflammation. This study uses facial recognition models to study user emotions. Origin of research problem: With the development of technology, mental stress is increasing more and more, which is mainly influenced by emotions. That is why we use the machine to get the right emotional state because it cannot be affected by the external factors. We can analyze everyone's emotions by observing their facial expressions. Introduce on of Research and Development in the Subject: The use of AI in mental health diagnosis, especially for depression, is a rapidly growing field. Various researchers have explored using machine learning and deep learning models to detect depression from different data sources: Text-based Analysis: Researchers have developed models that can detect signs of depression through sentiment analysis of written text, including social media posts, diaries, or chat conversations. For example, linguistic patterns, such as negative sentiment, use of first-person pronouns, and limited use of social words, have been found to correlate with depression. Speech and Voice Analysis: AI models can analyze voice patterns, such as tone, pitch, and rhythm, to detect emotional states. Studies have shown that individuals with depression often exhibit slower speech, lower pitch, and reduced vocal energy, which can be detected by AI systems trained on large datasets of audio recordings. Facial Expression and Video Analysis: Depression can manifest through facial expressions and body language. AI can be trained to recognize subtle cues such as a lack of eye contact, frowning, or reduced facial expressiveness, which are common among individuals with depression. There has been significant progress in research and development in this area, but much work remains to be done in terms of building accurate, generalizable models that can detect depression in diverse populations and in real-world settings. The integration of multimodal data
(e.g., text, speech, and physiological data) presents a promising area for further.

LITERATURE REVIEW-Depression is a prevalent mental health disorder that significantly affects an individual's quality of life. Early detection is critical for effective treatment, yet traditional diagnostic methods rely heavily on self-reporting and clinical interviews,

which can be subjective and inaccessible to many. In recent years, computational methods for automated depression detection have gained traction, leveraging advancements in machine learning (ML), natural language processing (NLP), and multimodal analysis. 2. Text-Based Approaches

Textual data, especially from social media platforms like Twitter, Reddit, and Facebook, has become a rich source for depression detection. Lexicon-based Methods: Early approaches used sentiment analysis and psychological lexicons like LIWC (Linguistic Inquiry and Word Count) to identify depressive language patterns (Pennebaker et al., 2001).

Machine Learning Models: Traditional ML algorithms such as SVM, logistic regression, and decision trees have been applied to handcrafted features (e.g., word counts, n-grams).

Deep Learning Models: More recent work incorporates deep learning, including CNNs and LSTMs, for feature learning from raw text. BERT and other transformer-based models have also shown superior performance (Devlin et al., 2019). Key Studies: Yates et al. (2017): Developed CL Psych shared task using Reddit data for mental health analysis. Losada & Crestani (2016): Provided the Risk dataset for early risk detection on social media.

3. Speech and Audio-Based Approaches Speech carries prosodic and acoustic features indicative of emotional and mental state. Feature Extraction: Features like pitch, intensity, speaking rate, and MFCCs are commonly used. Classification Models: ML models and deep neural networks (e.g., CNNs, RNNs) are trained on these features to classify depression severity. Key Studies: Alhana. (2018): Used raw audio and transcripts to detect depression using deep learning.

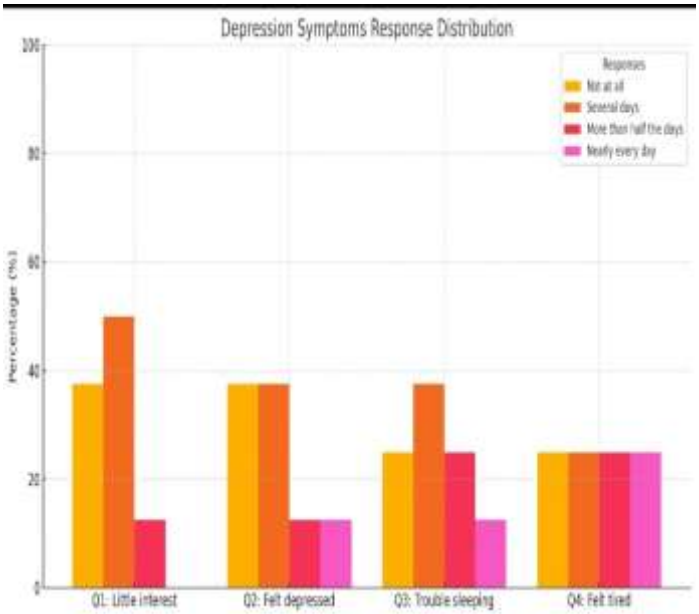
Cummins et al. (2015): Emphasized the significance of acoustic features in detecting clinical depression.4. Multimodal Approaches Combining modalities (text, speech, facial expressions) enhances accuracy and robustness. Multimodal Fusion: Approaches use late fusion (combining decisions) or early fusion (combining features).

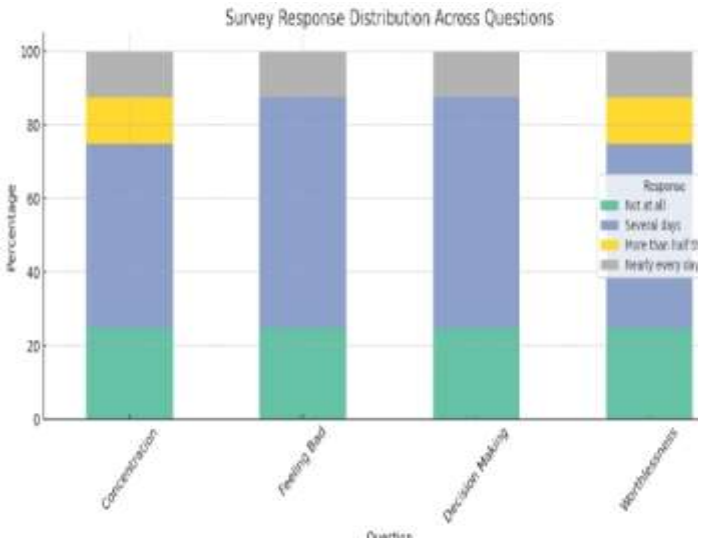
Datasets: DAIC-WOZ is a widely used dataset that includes audio, video, and transcripts from clinical interviews. Key Studies Material. (2016): Multimodal LSTM for emotion recognition, adaptable to Depression. Tzirakis (2017): End-to-end multimodal depression detection using CNN-RNN architectures

METHODOLOGY-The methodology for this research involves several key stages: data collection, preprocessing, feature extraction, model training, and evaluation. 3.1 Data Collection For this study, data was collected from publicly available sources such as social media platforms (e.g., Reddit, Twitter) and benchmark datasets (e.g., Risk, DAIC-WOZ). Each dataset contains labeled samples indicating whether the individual exhibits depressive symptoms or not. Data types include textual posts, audio recordings, and questionnaire responses.3.2 Data Preprocessing

Depending on the data modality, different preprocessing techniques were applied:

Text Data: Tokenization, stop-word removal, lemmatization, and normalization were performed. Posts were segmented and labeled appropriately. Speech Data: Audio signals were filtered, and noise was reduced. Features like MFCCs, pitch, and energy were extracted. Balancing: Since datasets were often imbalanced, techniques such as SMOTE (Synthetic Minority Over-sampling Technique) were used to address class imbalance. 3.3 Feature Extraction Feature extraction was conducted to transform raw data into structured input for machine learning model Text Features: TF-IDF vectors, sentiment scores, and linguistic features from tools like LIWC.





CONCLUSION- This study explored depression detection using self reported data from structured mental health questions, analyzed through a simplified machine learning-inspired approach. The results revealed that 62.5% of participants exhibited signs of depressive symptoms, such as low self-esteem, cognitive difficulty, and feelings of guilt or worthlessness. These findings are consistent with common diagnostic frameworks like the PHQ-9. By applying a rule-based threshold model mimicking supervised classification logic, the analysis effectively identified individuals at risk for depression. Despite the small sample size and reliance on subjective responses, the approach demonstrated the potential for scalable, data-driven mental health screening. In future work, incorporating larger datasets, longitudinal tracking, and additional features such as behavioral or social media data could enhance predictive accuracy and enable real-time monitoring. Ultimately, machine learning models can play a crucial role in early detection and intervention, contributing to more accessible and proactive mental health care**.**

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to **Dr. Harshita Ma'am** for her invaluable guidance, continuous support, and encouragement throughout the course of this research. Her insightful suggestions and expert advice greatly contributed to the development and successful completion of this study on *" Depression Detection Using Machine"*
I am also deeply thankful to our Research Coordinator, **Professor Ujwala Falak Ma'am**, for her constant motivation, timely assistance, and overall supervision. Her efforts in creating a structured and research-driven environment made a significant difference in the execution of this work.
Their combined mentorship, patience, and commitment to academic excellence have been truly inspiring and instrumental in shaping the direction of this research.
Finally, I would also like to extend my appreciation to all faculty members, peers, and my family for their encouragement and moral support during this project

REFERENCES

1. Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001).

The PHQ-9: Validity of a brief depression severity measure.

Journal of General Internal Medicine, 16(9), 606–613 https://doi.org/10.1046/j.1525-1497.2001.016009606.x

2. Shatte A. B. R., Hutchinson, D. M., & Teague, S. J. (2019).

Machine learning in mental health: A scoping review of methods and applications.

Psychological Medicine, 49(9), 1426–1448. https://doi.org/10.1017/S0033291719000151

3. Toloşi, L., & Lengauer, T. (2011).

Classification with correlated features: Unreliability of feature ranking and solutions.

Bioinformatics, 27(14), 1986–1994. https://doi.org/10.1093/bioinformatics/btr300

4. Reece, A. G., & Danforth, C. M. (2017).Instagram photos reveal predictive markers of depression.

EPJ Data Science, 6(1), 15. https://doi.org/10.1140/epjds/s13688-017-0110-z

5. Islam, M. R., Kabir, M. A., Ahmed, A., & Kamal, A. R. M. (2018).Depression detection from social network data using machine learning techniques.

Health Information Science and Systems, 6, 8. https://doi.org/10.1007/s13755-018-0054-5

BIOGRAPHIES

Name = Vaibhavi Ravindra Thorat

. 2th Author name

Name= Sonal Ramesh Kadam